A Proportion of Self-Collected Rectal Swabs Yield Human Immunodeficiency Virus Sequences Phylogenetically Related to Those from Plasma Human Immunodeficiency Virus RNA

Hannah Hudson,¹ Richard D'Aquila,¹ Brian Mustanski,^{2,3} and Ethan Morgan²

Abstract

We determined HIV-1 *pol* gene sequences from self-collected rectal swabs of HIV-positive young men who have sex with men and transgender women. HIV-1 *pol* was amplified from 39/96 (41%) rectal swabs, including 29/77 (38%) prevalent and 10/19 (53%) incident HIV-1 infections ($p < .001$). *Pol* did not amplify from rectal swabs from participants with plasma viral load $< 1,000$ copies/mL. Each rectal swab-derived amplicon consensus sequence was most closely related to the paired plasma virion RNA-derived sequence from the same participant. Results document a rectal mucosal source of HIV-1 in infected persons and suggest usefulness for noninvasive study of biological mechanisms underlying the epidemiologic risk to an insertive partner of HIV-1 acquisition during condomless anal sex.

Keywords: HIV, mucosal, phylogenetics

STRONG EPIDEMIOLOGIC EVIDENCE indicates that an insertive partner is at risk of HIV-1 acquisition during condomless anal sex.^{1–4} There is, however, limited evidence whether cells in the anorectal mucosa may be the source of HIV-1 transmitted from an infected receptive partner to an uninfected insertive partner.^{1–4} It is also not known if HIV-1 sequences in plasma RNA are phylogenetically related to those derived from cellular sources in the rectal mucosa of the same infected person. We studied whether HIV-1 *pol* gene sequences could be obtained from self-collected rectal swabs and compared them with sequences derived from HIV-1 plasma RNA, to begin addressing these issues. We studied both incident and prevalent HIV-seropositive participants in a cohort of young men who have sex with men and transgender women (YMSM/TW) in Chicago.

Participants were enrolled in RADAR, an ongoing longitudinal cohort study of health issues among diverse YMSM/TW living in the Chicago metropolitan area^{5,6}. A variety of complementary methods for participant recruitment were selected to achieve the multiple cohort, accelerated longitudinal design.⁷ At the time of enrollment into their original respective cohorts, all participants were between 16

and 29 years of age, assigned male at birth, spoke English, and had a sexual encounter with a man in the previous year or identified as gay, bisexual, or transgender. Finally, cohort members were allowed to refer a maximum of three peers for enrollment into the study as long as they were between 16 and 29 years of age. All cohort members continue to complete follow-up visits at 6-month intervals.

Participants perform self-administered rectal swabs to test for rectal *Neisseria gonorrhoeae* and *Chlamydia trachomatis* infections at each visit (Aptima Combo 2 Assay; Hologic, Toronto, CA). Specific instructions for atraumatically self-collecting the swab are described in the section “Rectal Swab Instructions” of Supplementary Data.

HIV diagnostic testing was done at all visits on those testing negative at a prior visit, using a point of care (POC) combination HIV antigen/antibody test from fingerstick blood (Alere Determine HIV 1/2 Ag/Ab POC; Abbott Diagnostics, Lake Forest, IL). Positives on the POC test are subsequently confirmed as HIV seropositive from venipuncture blood by Centers for Disease Control and Prevention-recommended laboratory test sequence of fourth generation antigen/antibody combination immunoassay,

¹Division of Infectious Diseases and HIV Translational Research Center, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois.

²Institute for Sexual and Gender Minority Health and Wellbeing, Northwestern University, Chicago, Illinois.

³Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, Illinois.

HIV1/2 antibody differentiation immunoassay, and, if needed, an HIV-1 nucleic acid test.⁸

Quantitative plasma HIV-1 RNA levels are determined using an assay with lower limit of detection of 40 copies/mL (Abbott RealTime HIV-1 Viral Load assay on the Abbott m2000 platform; Abbott Diagnostics). Plasma HIV-1 virion RNA is sequenced from samples with >40 copies/mL from entry to the study with preexisting HIV-1 infection or a later visit when seropositivity is first detected. HIV *pol* sequence generation utilizes reverse transcription (RT) and PCR followed by consensus amplicon Sanger sequencing of HIV-1 *pol* region (nucleotides 2,011–3,420 of HIV-1 HXB2 genome GenBank accession no. K03455; Genewiz, South Plainfield, NJ).

Self-collected rectal swabs (Aptima Unisex Swab Specimen Collection Kit for Female Endocervical and Male Urethral Swab Specimens; Hologic) are obtained from participants testing HIV positive on the POC assay coincident with collection of confirmatory HIV diagnostic testing, viral load, and plasma HIV sequencing. Two separate swabs are self-collected by participants during the study visit; one swab is used for rectal sexually transmitted infection testing and the other for sequencing. Either the first or second self-collected swab could have been used for sequencing as they were not labelled differently.

Rectal swabs used for sequencing were stored in 2.9 mL swab transport buffer (Aptima) at -20°C before analysis. In almost all cases, the entire amount of swab transport buffer was thawed at room temperature and processed as follows. Thawed transport buffer was transferred to a 15 mL conical tube and spun for 10 min at 1500 g to remove debris. The supernatant was transferred to a 5 mL ultracentrifuge tube (Beckman Coulter part #326819) and pelleted at 106,469 g for 90 min at 10°C . The supernatant was poured off, and the pellet was used for viral RNA extraction. Viral RNA was extracted and eluted into a final volume of 30 μL (QIAamp Viral RNA Mini Kit). The extracted viral RNA was then concentrated (Zymogen RNA Clean and Concentrator) and eluted into a final volume of 8 μL .

RT was performed according to the manufacturer's directions with an input of 8 μL of viral RNA (Invitrogen SuperScript III First-Strand Synthesis System). Nested PCR was performed (Invitrogen Platinum Taq DNA Polymerase High Fidelity). Two different sets of nested primers were used since some specimens amplified with only one set (see section "Sequencing Methods and Controls" of Supplementary Data for details of primers and amplification conditions). All PCR fragments were run on a gel to verify size before Sanger sequencing of amplicons (Genewiz). In only four cases, half the volume of swab transport buffer was refrozen without processing as above; the second half of the swab transport buffer was rethawed later for control experiments lacking RT or using DNase pretreatment. DNase treatment used Ambion AM1906 DNase Treatment and Removal Reagents (Thermo Fisher Scientific).

HIV *pol* amplification and sequencing were attempted from 96 self-collected rectal swabs paired with a plasma from the same visit. PCR amplification with at least one set of *pol* primers was successful from rectal swabs of 39/96 (41%) paired specimens, including 29/77 (38%) prevalent and 10/19 (53%) incident cases ($p < .001$); see Supplementary Data for numbers amplified with each primer pair. Sequences were determined from every amplicon obtained. Plasma viral loads

for those participants with swabs yielding amplicon sequences ranged from 1,565 to 2,558,478 copies/mL. Fifty-five percent of swabs from those with plasma viral load >100,000 copies/mL yielded sequence, 48% of those with plasma viral load >10,000 copies/mL yielded sequence, and 41% of those >1,000 copies/mL yielded sequence. HIV *pol* was not amplified from any of the 10 rectal swabs from participants with plasma viral loads ranging from 689 to 118 copies/mL. A significantly greater proportion of incident cases (10/19, 53%) had a viral load >1,000 copies/mL, compared to prevalent cases (29/77, 38%; $p < .001$).

Rectal swab-derived amplicon sequences were compared to plasma-derived amplicon sequences obtained at the same visit from 29 prevalent and 10 incident cases. Each was aligned to the Consensus B HIV reference sequence using MUSCLE (Multiple Sequence Comparison by Log-Expectation, European Bioinformatics Institute) multiple sequence alignment⁹ in the MEGA v7.0 (Molecular Evolutionary Genetics Analysis) software package.¹⁰ Consensus bulk amplicon sequences were used; minor variants or ambiguities were not considered. Phylogenetic tree analyses were performed using the neighbor-joining method¹¹ with pairwise distance calculated by TN93¹² analysis.

A phylogenetic tree was rooted with HIV-1 consensus clade B sequence along with 200 clade B sequences from men in the United States obtained from 2010 to 2015 (accession numbers in Los Alamos HIV Database provided in Supplementary Table S1). The tree indicates that each rectal swab-derived sequence was closest to the plasma RNA-derived sequence from the same participant (Fig. 1). Short branch lengths were found between rectal swab- and plasma RNA-derived sequences (Fig. 2). Median pairwise distance between rectal swab- and plasma-obtained viral sequences was 0.01 nucleotide substitutions per site (interquartile range = 0.023).

The first half volume of swab transport buffer from four specimens collected near the end of the study that had high plasma viral loads each yielded rectal swab-derived amplicons and sequences. In addition, the refrozen second half was thawed a second time to study if viral DNA or RNA was amplifiable. Half of the extracted HIV-1 nucleic acid from these four rethawed swab specimens was amplified without *in vitro* RT. Three of the four did not amplify, consistent with prior amplification from an RNA source in those three specimens. The remaining half of the extracted nucleic acid from the single rectal specimen that amplified without *in vitro* RT was then split into two conditions: DNase treatment versus no DNase treatment. RT was then performed on each. Amplification was subsequently obtained both with and without prior DNase treatment. Together, these results suggest amplification from DNA, as well as RNA, from that single rectal swab. This limited sample size suggests that the type of nucleic acid amplified from rectal swabs can be determined, and more specimens must be studied to draw any conclusions.

We obtained HIV-1 *pol* amplicon sequences from about half of rectal swabs self-collected by YMSM/TW with plasma viral load >1,000 copies/mL. More work is needed to assess whether amplicon recovery was limited by the swab type, the self-collection not always obtaining adequate sample, antiretroviral therapy that incompletely suppressed viremia, or biological variability. The close genetic

relatedness of sequences from each participants' rectal swab and plasma excludes possibilities that rectal swab-derived sequences were either laboratory contaminants or from semen remaining after recent condomless anal sex. Results also raise questions not addressed here about whether rectal swab-derived HIV-1 sequences are more or less "compartmentalized" than those at other tissue sites, and whether viruses "exchange" in either direction between rectal tissues and plasma.

Even if rectal swab sequences are determined from only a subset of those studied, this offers an alternative noninvasive approach that can complement colon biopsies to characterize HIV-1 in these tissues. It can potentially help inform about which of the cellular sources of HIV-1 in deeper colon mucosa produce the viruses that reach the epithelial surface for transmission to a condomless insertive partner and help begin characterizing mechanisms underlying the transmission risk of insertive condomless anal sex.¹⁻⁴ This could potentially also be applied to simian immunodeficiency virus-macaque models, as well as other mucosal surfaces (e.g., female reproductive tract or oral mucosa). In addition, the use of rectal swabs may have some utility in molecular HIV surveillance for public health purposes^{13,14} in settings where either venipuncture or dried blood spot collection may be impractical, if amplicon yield can be improved.

Acknowledgments

The authors thank the entire RADAR research team, particularly Dr. Thomas Remble and Antonia Clifford for overseeing the project and Daniel T. Ryan for data management. The authors also thank the RADAR participants.

Author Disclosure Statement

No competing financial interests exist.

Funding Information

This work was supported by grants from the National Institute on Drug Abuse (U01DA036939, PI: B.M.; F32DA046313, PI: E.M.) at the National Institutes of Health (NIH). This work also benefited from collaborative infrastructure enabled by the Third Coast Center for AIDS Research (CFAR), an NIH funded center (P30 AI117943).

Supplementary Material

Supplementary Data
Supplementary Table S1

References

1. Baggaley RF, Owen BN, Silhol R, *et al.*: Does per-act HIV-1 transmission risk through anal sex vary by gender? An updated systematic review and meta-analysis. *Am J Reprod Immunol* 2018;80:e13039.

2. Baggaley RF, White RG, Boily MC: Infectiousness of HIV-infected homosexual men in the era of highly active antiretroviral therapy. *AIDS* 2010;24:2418–2420.
3. Jin F, Jansson J, Law M, *et al.*: Per-contact probability of HIV transmission in homosexual men in Sydney in the era of HAART. *AIDS* 2010;24:907–913.
4. Vittinghoff E, Douglas J, Judson F, McKirnan D, MacQueen K, Buchbinder SP: Per-contact risk of human immunodeficiency virus transmission between male sexual partners. *Am J Epidemiol* 1999;150:306–311.
5. Johnson BT, Redding CA, DiClemente RJ, *et al.*: A network-individual-resource model for HIV prevention. *AIDS Behav* 2010;14(Suppl 2):204–221.
6. Mustanski B, Garofalo R, Herrick A, Donenberg G: Psychosocial health problems increase risk for HIV among urban young men who have sex with men: Preliminary evidence of a syndemic in need of attention. *Ann Behav Med* 2007;34:37–45.
7. Miyazaki Y, Raudenbush SW: Tests for linkage of multiple cohorts in an accelerated longitudinal design. *Psychol Methods* 2000;5:44–63.
8. DiNenno EA, Prejean J, Irwin K, *et al.*: Recommendations for HIV screening of gay, bisexual, and other men who have sex with men—United States, 2017. *MMWR Morb Mortal Wkly Rep* 2017;66:830–832.
9. Edgar RC: MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;32:1792–1797.
10. Tamura K, Stecher G, Peterson D: MEGA7: Molecular evolutionary genetic analysis version 7.0. *Mol Biol Evol* 2013;30:2725–2729.
11. Saitou N, Nei M: The neighbor-joining method—a new method for reconstructing phylogenetic trees. *Jpn J Genet* 1986;61:611.
12. Tamura K, Nei M: Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 1993;10:512–526.
13. Wertheim JO, Leigh Brown AJ, Hepler NL, *et al.*: The global transmission network of HIV-1. *J Infect Dis* 2014; 209:304–313.
14. Centers for Disease Control and Prevention: Detecting, investigating, and responding to HIV transmission clusters. Health and Human Services, Atlanta, GA, 2017.

Address correspondence to:

Brian Mustanski
Institute for Sexual and Gender Minority Health
and Wellbeing
Northwestern University
625 North Michigan Avenue, Suite 14-61
Chicago, IL 60611

E-mail: brian@northwestern.edu